Euphytica (2010) 171:417–425
DOI 10.1007/s10681-009-0072-8

Response to phenotypic and marker-assisted selection for yield and quality component traits in cucumber (*Cucumis sativus* L.)

T. K. Behera · Jack E. Staub · Snigdha Behera · Shanna Mason

Received: 20 April 2009 / Accepted: 28 October 2009 / Published online: 7 November 2009

Abstract Two cucumber recombinant inbred lines (RILs) differing in plant habit were crossed and progeny self-pollinated to produce F_3 individuals upon which phenotypic selection was practiced to identify a base population which in turn underwent either two cycles of MAS or random mating without selection (RAN). MAS and RAN were practiced to produce F_4 and F_5 progeny sets. RIL, crossing parents, and F_3–F_5 progeny sets were then evaluated under replicated field conditions for fruit yield and quality (L:D and E:T) to evaluate gain from selection (ΔG). The broad-sense heritability (h^2B) over cycles (C) of selection ranged 0.22–0.45, 0.09–0.20, and 0.11–0.15 for yield, L:D, and E:T, respectively. Although one cycle of PHE selection followed by MAS was effective in conserving the performance of the traits examined during inbreeding, progeny performance during RAN fluctuated (F_4–F_5 generation; C_2). Lack of ΔG during advanced generations (F_4–F_5) of MAS was likely due to allelic fixation and/or optimized epistatic complementation.

T. K. Behera (✉)
Division of Vegetable Science, Indian Agricultural Research Institute, PUSA Campus, New Delhi 110 012, India
e-mail: tusar@rediffmail.com

J. E. Staub · S. Behera · S. Mason
USDA, ARS, Vegetable Crops Research Unit, Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison, WI 53706, USA

Keywords Genetic markers · MAS · QTL analysis · Multi-trait selection · Genetic gain

Introduction

Molecular genotyping can enhance gain from selection (ΔG) during plant improvement when compared to phenotypic selection (PHE; Kasha 1999; Ortiz 1998; Fan et al. 2006), especially if marker-trait associations are robust and environment × genotype interactions have been characterized and successfully employed in PHE selection strategies (Staub et al. 1996; Paterson et al. 1991). The successful application of marker-assisted selection (MAS) in plant improvement will, however, ultimately depend upon its increased resource allocation efficacy (i.e., labor and cost) when compared to PHE selection (Xie and Xu 1998). The efficiency of MAS depends on several factors including marker number and kind (codominant vs. dominant), the strength of marker associations with selection indices, population size, and trait heritability (Gimelfarb and Lande 1994). The selection efficiency of MAS can, in fact, be superior to PHE selection even when target trait heritability is relatively low and population size is rather small (<50 individuals; Moreau et al. 2004; Fan et al. 2006).

The appropriate application of selection indices and marker-trait data can enhance ΔG (Lande and

Thompson 1990; Fan et al. 2006). Indeed, marker genotyping can improve trait introgression during MAS backcrossing and augment population development during conventional PHE selection by reducing the expense of tedious trait-based selection (Tanksley et al. 1981; Edwards et al. 1987; Lande and Thompson 1990; Dudley 1993; Knapp 1998). Phenotypic changes during MAS are typically associated with shifts in allelic frequency at loci linked to economically important traits under selection (Steele et al. 2004; Flint-Garcia et al. 2003; Moreau et al. 2004; Fan et al. 2006). Thus, introgression of such desirable alleles during marker-assisted backcrossing has predictably proven effective in several crop species (Willcox et al. 2002; Lecomte et al. 2004; Thabuis et al. 2004; Fan et al. 2006).

Yield in cucumber (*Cucumis sativus* L.) is conditioned by quantitative trait loci (QTL) whose affects are modulated by epistatic and genotype × environment interactions (Serquen et al. 1997a, b; Fazio et al. 2003a). Restriction fragment length polymorphisms (RFLPs; Kennard and Havey 1995), random amplified polymorphic DNAs (RAPDs; Serquen et al. 1997a), simple sequence repeats (SSRs; Fazio 2001), sequence characterized amplified regions (SCARs; Fazio 2003a), and single nucleotide polymorphisms (SNPs; Robbins 2006) have been used to identify the genomic position of yield and quality component QTL on moderately saturated linkage cucumber maps (Serquen et al. 1997a; Fazio et al. 2003a). Such marker-trait relationships have proven useful in MAS mediated backcrossing for yield components (Fazio et al. 2003a; Fan et al. 2006). However, PHE selection for disease resistance and sex expression in cucumber can negatively affect fruit yield and quality (Staub and Grumet 1993; Staub et al. 1986), and furthermore various yield and quality component traits are negatively correlated (Serquen et al. 1997a, b; Fazio et al. 2003a). Consequently, comparative assessments of marker-assisted and PHE selection strategies in relation to correlated responses to selection are critical for the successful implementation of MAS in cucumber improvement. Thus, a study was designed to: (1) Compare population response to MAS after initial PHE selection for improved fruit yield and quality during line development (inbreeding), and; (2) define allelic frequency changes during such selection. This information will assist in the construction of inbred lines prior to hybrid development where MAS and/or a combination of MAS and PHE selection might be employed.

Materials and methods

Parents

Two recombinant inbred lines (RIL) 7026B76 (B76) and 7022C8 (C8) were used as parents for population development. They were originally derived from the mapping population involving G7 and H19 in their pedigree. These two RILs with contrasting phenotypes (earliness, sex expression, branching, and fruit length:diameter ratio) were drawn from the U.S. Department of Agriculture (USDA) cucumber breeding program (Fazio et al. 2003a). Based on open-field evaluation, these monoecious parental lines contrasted significantly ($P < 0.05$) for the relative days (d) to anthesis from transplant [B76 (43d) vs. C8 (39d)], average number of lateral branches (NLB) in the first four flowering nodes [B76 (2) vs. C8 (2.6)], average number of pistillate flowers in the main stem over the first 10 nodes above the cotyledon [B76 (1.8) vs. C8 (1.2)], and seed cavity diameter [endocarp (E): total (T) diameter] ratio [B76 (5.8) vs. C8 (6.1)] (Fazio 2001).

Development and evaluation of F_3 progeny

The F_1 progeny (B76 × C8) were pollinated to produce 100 F_2 individuals that were then self-pollinated to produce 100 F_3 families [designated as the phenotypic (PHE) base population (BASE)]. In the summer of 2006 and 2007, parents, F_1 progeny, and F_3 families were sown as direct seed and evaluated (below) for four-harvest yield (fruit number), fruit length and diameter ratio (L:D), and E:T at the University of Wisconsin Central Sands Experiment Station (UWESH), Hancock, Wisc. [Soil type: Planefield loamy sand (Typic Udipsamment)]. Plots were arranged in randomized complete block design with three replications per location. Each replication had 12 plants and consisted of single rows with plants spaced 13 cm apart in rows to include edge borders positioned on 1.5 m centers corresponding to a plant density of ~51,000 plants/ha.

The diameter of 5–10 randomly selected fruits (USDA 2B–3A grade; 25–30 mm in diameter) in each experimental plot at each of four harvests was measured and used to calculate a cumulative four-harvest mean L:D and E:T ratio for each entry. Likewise, the number of fruit per plot was recorded for each harvest, and these data were used to

calculate the cumulative four-harvest yield per entry. The first harvest interval of each plot was determined when two to three fruit >51 mm in diameter (oversized) were observed within a plot (Wehner 1989). The remaining three harvest intervals occurred every 6–7 days when 2–3 mature oversized fruits were observed within a plot. All immature fruits >20 mm in diameter and >10 cm in length (USDA 1A–3B grade) were harvested.

Test family development

Based on phenotypic evaluation in 2006 (above), 16 of 100 F_3 families (designated individually by hyphenation; e.g., F_3-1) were selected for use as test arrays (designated PHE arrays) for the development of advance generations (F_4–F_5) by marker-assisted selection (MAS) using QTL associated with yield (fruit number) and quality (L:D and E:T ratios). Four test arrays were designated as low yield (<55 fruits/harvest) and low quality (i.e., low L:D and high E:T; e.g., F_3-12, F_3-39, F_3-82, and F_3-55), low yield and high quality (e.g., F_3-13, F_3-91, F_3-117, and F_3-32), high yield (>56 fruits/harvest) and low quality (e.g., F_3-70, F_3-86, F_3-100, and F_3-89), and high yield and high quality (2.8–3.2 L:D and 0.60–0.62 E:T***) (e.g., F_3-20, F_3-105, F_3-69, and F_3-60). Six plants of each of these 16 (4 × 4) F_3 families (total 96 plants) were grown in a greenhouse in Madison Wisc. for self-pollination and genotyping (12 DNA markers; Table 1) for use in MAS. Twenty F_4 test families were then chosen for test array development based on genotype (yield and quality associated QTL) and seed amounts required for replicated trialing (>60 seeds) (designated MAS arrays). Subsequently, four genotyped plants from each F_4 family (4 × 20 = 80 plants) that had been selected for QTL to form test array groupings (4) were used to produce 60 F_5 families (3 × 60 = 180 plants) whose genotype was confirmed by DNA analysis. Only those families homozygous at marker loci for associated traits (e.g., yield, L:D, or E:T) in each test array were used in the final phenotypic open-field evaluation [i.e., PHE (BASE), MAS, and RAN arrays].

Marker genotyping of test families

The individuals of each test family used herein (F_3–F_5) were genotyped using 12 mapped markers (four SNPs, five SCARs, two simple SSRs, and one RAPD; Serquen et al. 1997b; Fazio et al. 2003a) (Table 1) according to Fazio et al. (2003a). These loci were

Table 1 Characteristics of molecular markers in cucumber (*Cucumis sativus* L.) used for marker-assisted selection

Marker	Marker type	Linkage group	Map position (cM)	Parent[a]	Ideotype	QTL associations (mapping parent and LOD score)
AT1SNPG3H3[d]	SNP	6	63.8	G&H	G	Robbins (2006)
L1LG3H3[d]	SNP	6	115.00	G&H	G	Robbins (2006)
M8SNPG3H[d]	SNP	6	39.1	G&H	H	Robbins (2006)
W7SNPG1H3[d]	SNP	1	34.91	G&H	G	Robbins (2006)
AK5SCAR[b]	SCAR	6	33.00	G	H	MLB (H, 3.0) (Fazio et al. 2003b)
AW14 SCAR[b]	SCAR	3	3.89	G&H	G	GYN (G, 5.1) (Fazio et al. 2003a)
BC231SCAR[d]	SCAR	7	28.13	H	H	Fazio (2001)
BC523SCAR[b]	SCAR	1	48.79	G	H	MLB (H, 3.3) (Fazio et al. 2003b)
P14 SCAR[d]	SCAR	1	49.87	H	G	Fazio (2001)
CSWTAA0B[c]	SSR	3	24.60	G	H	Smaller E:T (G,8.5) (Behera et al. 2008)
CSWTAAA01[b]	SSR	4	34.10	G&H	H	MLB (H, 4.6) (Fazio et al. 2003b)
OPAT15-1[c]	RAPD	1	34.04	G	H	Smaller E:T (G, 6.6) (Behera et al. 2008)

[a] Allelic constitution based on mapping parents H-19 and Gy-7 (Fazio et al. 2003b), where *G* present in Gy421, *H* present in H-19, *G&H* present in Gy421 and H-19 (codominant marker)

[b] Markers associated with yield QTLs like *GYN* gynoecious, *MLB* multiple lateral branching, and quality QTLs (S:D ratio and fewer seeds)

[c] Markers associated with fruit quality traits (E:T) (Behera et al. 2008)

[d] Arbitrary primers were taken for assessing the diversity within the populations

polymorphic in the parental lines, and are associated with yield (e.g., fruit number) and quality (e.g., L:D; E:T) components (Fazio et al. 2003a; Robbins 2006; Behera et al. 2008). These marker loci (two alleles/locus) have been used successfully in MAS (Fazio et al. 2003b; Fan et al. 2006), and, thus, were used herein to develop germplasm test arrays (i.e., MAS) for analysis. All marker map locations were reported by Fazio et al. (2003a), except marker loci SSR-CSWTAA0B and RAPD-OPAT15-1 associated with quality (E:T) have been mapped on Linkage Group 3 and 1, respectively, which were described by Behera et al. (2008).

Open field evaluation

The original parental mapping lines (Gy-7, H-19; Fazio et al. 2003a), derived F_8 lines B76, and C8, and their resultant 100 F_3, 20 F_4, and 60 F_5 families, as well as 'Vlasset' (Seminis Seed Company, Woodland, Calif.; control) were sown on May 18, 2007 in a greenhouse in Madison, Wisc., and transplanted on June 12 at UWESH. The design was a randomized completed block design with three replications. Experimental plots consisted of ten plants spaced 13 cm apart within rows (5.2 m long) on 1.5 m row centers ($\sim$51,000 plants/ha) with end and side borders of 'Vlasset'. Data were collected on yield (number of fruit number), and L:D and E:T ratios as in the 2006 F_3 evaluation as described above.

Statistical analyses

Trait data were subjected to analyses of variances using a mixed models procedure (PROC Mixed) in SAS to determine treatment effects (Littell et al. 1996). Treatments of populations and cycles were considered fixed effects, while blocks were considered random. Least square mean comparisons of mean trait values were then performed using SAS (SAS Institute 1999). Broad-sense heritability (h^2B), and genetic (σ^2G) and environmental (σ^2E) variances were calculated using SPAR 1 computer software [Indian Agricultural Statistics Research Institute (IASRI), New Delhi, India]. To determine trait relationships, pairwise phenotypic Pearson correlations were calculated using SAS (SAS Institute 1999). Genetic gain from selection was measured in the F_4 and F_5 generations as departures from the BASE population (F_3) as $\Delta G = ih^2B\sigma P$; where i is the selection differential, h^2B is broad-sense heritability, and σP is the phenotypic standard deviation.

Genetic changes in population structure were assessed using marker allele frequencies in F_3 (96), F_4 (80), and F_5 (180) progenies calculated as a proportion of individuals examined. Alleles were designated as "favorable" or "unfavorable" according to their association with the phenotypic trait performance of the parents of the RIL population [line Gy-7 (allelic designation G) and line H-19 (allelic designation H)] (Fazio et al. 2003a, b; Fan et al. 2006). For the traits examined herein, alleles in line B76 and line C8 were designated as G and H, respectively, to compare allelic frequency changes in terms of ΔG (Table 1).

Parental lines and resulting cross progeny were strategically intermated to create the optimal ideotype for the traits under selection [i.e., homozygous for the appropriate favorable alleles (G and H)]. Based on allelic differences at marker loci [SNP (AT1SNPG3H3, L1LG3H3, M8SNPG3H1, and W7SNPG 1H3), SCAR (AK5SCAR, AW14SCAR, BC231S CAR, BC526SCAR, and P14 SCAR), SSRs (CSWTAA0B and CSWTAAA01), and RAPD (OPAT15-1)], gene frequency, percent polymorphic loci, and mean heterozygosity were estimated using computer algorithms in "Tools for Population Genetic Analyses" (TFPGA) ver. 1.3 (Miller 1997). Estimates of genetic distance (GD; Nei 1972) were calculated using algorithms in POPGENE version 1.32 (Yeh and Boyle 1997). The direction and magnitude of allelic frequency changes between generations were characterized by assessing mean marker-associated trait differences over selection cycles [PHE + MAS (F_3–F_4) and RAN (F_4–F_5)] using regression analysis (Steele and Torrie 1980). Best-fit models (linear or quadratic) were identified and are presented herein based on comparative generation analyses.

Results

Analysis (ANOVA) detected highly significant ($P < 0.01$) differences between parental lines for yield (P_1; 18.6 $\pm$ 3.8 vs. P_2 16.1 $\pm$ 3.2), L:D (P_1; 3.00 $\pm$ 0.2 vs. P_2; 2.9 $\pm$ 0.2), and E:T (P_1; 0.60 $\pm$ 0.02 vs. P_2; 0.65 $\pm$ 0.03), and between parental lines and cross-progeny (grand mean = 16.9 $\pm$ 6.98,

2.9 ± 0.79 and 0.62 ± 0.03, respectively), and among F_3, F_4, and F_5 progenies for these traits in 2007 (data not presented). Means and ranges of F_3, F_4, and F_5 populations, as well as variance components and broad-sense heritability estimates are presented in Table 2. Differences in mean values among segregating F_3, F_4, and F_5 progeny populations were remarkable, ranging from 10.5 to 24.8 for cumulative three-harvest yield (fruits/plant), 2.6–3.3 for L:D, and 0.56–0.67 for E:T. Similarly, h^2B estimates ranged from 0.22 to 0.45 for yield, 0.09–0.20 for L:D, and 0.11–0.15 for E:T. The ΔG for yield as a function of performance compared to the BASE population (F_3) for F_4 and F_5 progeny was a +6.9 and +6.7 units, respectively. Nevertheless, the environmental variance associated with yield was relatively high 53.5–73.0% across generations when compared to L:D (0.55–1.16%) and E:T (0.001–0.003%).

Trait correlations

Positive phenotypic correlations between the traits examined were low [0.08–0.17; e.g., yield vs. E:T = 0.08 in F_5] in all generations (F_3, F_4 and F_5) (Table 3), whereas the low negative correlations (−0.06 to −0.17 for F_3–F_5 progeny) were observed between yield and L:D. In contrast, a significant negative correlation (−0.46, $P = 0.05$) was detected between L:D and E:T during one cycle of PHE + MAS (F_3–F_4). Likewise, a significant negative correlation (−0.39, $P = 0.05$) was detected between these traits during one cycle (F_4−F_5) of RAN.

Table 3 Correlation coefficients among three different traits across the F_3, F_4, and F_5 generations during one cycle of phenotypic and marker-assisted selection (MAS) in cucumber (*Cucumis sativus* L.)

Population[a]	Trait	L:D[c]	E:T[d]
F3	Yield[b]	−0.17 (0.9)	0.16 (0.10)
	L:D		−0.25 (0.01)
F4	Yield	−0.06(0.8)	0.17 (0.46)
	L:D		−0.46 (0.04)
F5	Yield	−0.16 (0.9)	0.08 (0.53)
	L:D		−0.39 (<0.01)

[a] F_3, F_4, and F_5 populations were generated from a cross between B-76 (B7) × C8 which originate from lines H-19 (H1) and Gy-7 (G7) and selected for yield and quality trait-associated quantitative trait loci

[b] Yield represented by number of fruit per plot was collected for each harvest, and is presented as cumulative four-harvest yield per entry

[c] Five to ten randomly selected fruits (USDA 2B-3A grade; 25–30 mm in diameter) in each experimental plot for each harvest were used to calculate a four-harvest mean L:D ratio (number in the parentheses represent the *P* values)

[d] Five to ten randomly selected fruits (USDA 2B-3A grade; 25–30 mm in diameter) in each experimental plot for each harvest were used to calculate a four-harvest mean E:T ratio (number in the parentheses represent the *P* values)

Marker frequency changes during MAS

Marker allele frequency changes were detected after PHE + MAS at some loci associated with yield and quality components (Table 4). For instance, allelic frequency changes (both increases and decreases) occurred at loci associated with the yield components,

Table 2 Means, ranges, genetic variances (σ^2G), environmental variances (σ^2E), and heritability (h^2B), and genetic gain ΔG for yield (number of fruits per plot per harvest), and fruit length (L):diameter (D) ratio and seed cavity diameter ratio [endocarp (E):total (T) diameter] in the F_3, F_4 and F_5 generations derived from a mating between cucumber (*Cucumis sativus* L.) inbred lines 7026B76 (B76) and 7022C8 (C8)

Trait	F_3			F_4			F_5		
	Yield	L:D ratio	E:T ratio	Yield	L:D ratio	E:T ratio	Yield	L:D ratio	E:T ratio
Mean	17.5	2.9	0.62	17.3	2.9	0.62	15.9	2.9	0.62
Range	(12.5–23.1)	(2.7–3.3)	(0.57–0.67)	(13.9–24.8)	(2.6–3.1)	(0.56–0.66)	(10.5–22.3)	(2.6–3.3)	(0.57–0.67)
SE	6.98	0.79	0.03	5.97	0.61	0.04	6.40	0.88	0.23
CV (%)	48.80	37.87	6.08	43.07	29.90	6.55	59.09	52.67	6.27
σ^2G	20.63	0.15	0.002	44.06	0.26	0.003	43.84	0.29	0.01
σ^2E	73.06	0.93	0.002	53.51	0.55	0.001	61.55	1.16	0.003
h^2B	0.22	0.14	0.12	0.45	0.09	0.15	0.42	0.20	0.11
ΔG	–	–	–	6.90	0.011	0.011	6.70	0.37	0.053

multiple lateral branching (MLB; AK5SCAR, CSWTAAA01, and BC523SCAR) and gynoecy (GYN; AW14SCAR). Nevertheless, although positive changes (i.e., those that promote positive ΔG; Fan et al. 2006) in allelic frequencies at QTL for GYN and MLB occurred across generations [e.g., AW14SCAR (0.84–1.00) and BC523SCAR (0.37–0.00); Table 4], these alterations were inconsistent with predicted associated phenotypic changes (e.g., three-harvest yield as mean fruits/plant for F_3 and F_5 progenies = 17.5 and 15.9, respectively) (Table 2). Given these generational fluctuations, appraisal and presentation of allelic frequency changes at the loci examined by regression analysis was deemed inappropriate.

Marker loci SSR-CSWTAA0B and RAPD-OPAT15-1 associated with quality (E:T ratio) have been mapped on linkage groups (LG) LG 3 and LG 1, respectively (Behera et al. 2008). The allele (H) at CSWTAA0B associated with smaller E:T changed from 0.79 to 1.00 (Table 4) and allelic frequencies at OPAT15-1 associated with smaller E:T also changed from 0.66 to 1.00 PHE + MAS (F_3–F_5) (Table 4). There was a positive response to PHE and MAS for E:T (i.e., F_3–F_5) associated with CSWTAA0B and OPAT15-1, and the E:T ratio across the same generations was observed to be 0.62 (Table 2).

Table 4 Allele frequencies at molecular marker loci associated with yield and quality components in cucumber (*Cucumis sativus* L.) following one cycle of phenotype selection (F_3–F_4) followed by marker-assisted selection 1 (F_4–F_5)

Marker[a]	Phenotypic association[b]	Expected frequency[c]	Family[d]		
			F_3	F_4	F_5
AK5SCAR	MLB	0.00	0.20	0.00	0.00
AW 14 SCAR	GYN	1.00	0.84	1.00	1.00
BC523SCAR	MLB	0.00	0.37	0.00	0.00
CSWTAAA01	MLB	0.00	0.33	0.00	0.00
CSWTAA0B	E:T	1.00	0.79	1.00	1.00
OPAT15-1	E:T	1.00	0.66	1.00	1.00

[a] Marker type, map location, and association with yield and quality components given in Table 1

[b] *MLB* multiple later branching, *GYN* gynoecy, and *E:T* = seed cavity diameter ratio [endocarp (E):total (T) diameter]

[c] Expected frequencies of the Gy-421 marker phenotype (G) based on marker QTL associations (Fazio et al. 2003a; unpublished data)

[d] F_3, F_4, and F_5 populations generated from a cross between B-76 (B7) × C8 which originate from lines H-19 (H1) and Gy-7 (G7) and were selected for yield and quality trait-associated quantitative trait loci

Population structure changes during MAS

Population genotyping allowed for the characterization of population changes in response to PHE and MAS. Generally, allelic frequency changes during PHE in F_3 families followed by MAS resulted in ΔG for yield and L:D after one cycle of selection for the traits examined. Individuals in the F_3 generation possessed the highest-level of polymorphism (92.8%), heterozygosity (0.34), and the greatest genetic diversity (GD = 0. 36) when compared to the other populations examined [F_4: polymorphism (53.2%, heterozygosity (0.19) and GD (0.19) and F_5: polymorphism (89.3%), heterozygosity (0.26) and GD (0.28)] (data not presented). Genetic affinities were detected between F_3 progenies and the parental line B76 the line from which it was derived (i.e., RIL parent H-19). These germplasms possessed QTL alleles associated with high yield potential. Similarly, genetic affinities were detected between F_3 progenies and F_5 resulting from one cycle of MAS for QTL associated with improved fruit yield and quality, and also found in parental H-19 and B76.

Diversity among progeny within a generation decreased [polymorphism = 53.2% and heterozygosity = 0.19] after one cycle of PHE + MAS (from F_3 to F_4; data not presented). Nevertheless, a positive response to selection was detected for yield and quality. In contrast, such generational diversity increased [polymorphism = 89.3% and heterozygosity = 0.26] after one cycle of RAN (from F_4 to F_5; data not presented). The GD among F_4 (0.19) and F_5 (0.28) progenies, however, was dramatically less than the GD between their parental lines (B76 and C8; GD = 0.37) and the RIL parents from which the parental lines were derived (Gy-7 and H-19; GD = 0.90) (data not presented; Fazio et al. 2003a).

Discussion

Improvement of fruit yield and quality is a major focus of many cucumber improvement programs (Lower and Edwards 1986, Tatlioglu 1993). MAS has been proven effective and efficient for increasing yield and quality in cucumber during backcrossing

(Fan et al. 2006; Fazio et al. 2003b). However, ΔG can vary depending on marker and selection type (e.g., backcross PHE vs. MAS; Fazio et al. 2003b; Robbins and Staub 2009). The time required to complete one cycle of open-field PHE cucumber selection in north temperate regions of the U.S. is one year (Fan et al. 2006). The time required for MAS and increase (F_3–F_5) of the three populations employed herein required only seven months. Thus, in cucumber, MAS could be cost-effective for improving overall selection efficiency for yield and quality depending on the traits examined and their genetic nature (gene number and action, linkage associations, heritability, etc.).

Yield in cucumber is quantitatively inherited, has a low heritability [i.e., narrow-sense heritability of 0.07–0.25], and is influenced by environment and to a lesser degree by genotype × environment interactions (Wehner 1989). Heritability (h^2B) for yield as estimated herein (F_3–F_5 family performance) was relatively low (0.22–0.45) across cycles of selection. These estimates are similar to those of Serquen et al. (F_3; 1997b) and Fazio et al. (RIL; 2003a), where cross-progeny were derived from the same genetic background (i.e., Gy-7 × H-19). Moreover, the heritability estimates for fruit L:D ratio calculated herein ($h^2B = 0.09$–0.20) were also low and similar to those of Fan et al. ($h^2B = 0.09$–0.11; 2006) who also used cross-progeny originating from Gy-7 and H-19.

Processing cucumbers in the U.S. are graded based on their size, with the smaller fruit usually bringing a higher price (Tatlioglu 1993). Thus, fruit dimensions (e.g., L:D) are considered yield components, since they determine marketable yield. For example, U.S. processing cucumbers must have an L:D between 2.9 and 3.3 to be commercially acceptable (Staub and Bacher 1997). After one cycle (PHE + MAS) of selection, significant ($P = 0.05$) but low correlations were detected between yield and L:D ratio (−0.06 to −0.17; Table 3). Given previous reports of such correlations between yield and L:D (Serquen et al. 1997a, $r = -0.98$, Fazio 2001, $r = -0.27$ to -0.36), the results presented herein corroborate the contention that as L:D increases, fruit yield may decrease in this genetic background (Gy-7 × H-19).

Genetic affinities among parental lines and their derived cross-progeny (F_3–F_5 families) were apparent. Genetic diversity in all cross-progeny populations was relatively high (data not presented), and GD between some of these populations (i.e., possessing relatively high fruit yield and moderate quality) and parental line Gy-7 (i.e., possessing alleles for low yield and quality) were remarkable. In one case, the allelic constitution of cross-progeny (F_3 vs. F_4) after one cycle of MAS for yield and quality provided a positive selection response. However, this was not maintained in the subsequent cycle of MAS (from F_4 to F_5). For instance, comparatively large genetic gains for yield were achieved [F_3–F_4 (6.9 units) and F_4–F_5 (6.7); Table 2] beyond the base population (F_3), but the phenotypic values in terms of yield were reduced in advanced generations (F_3-17.5; F_4-17.3 and F_5-15.9; Table 2). This may be indicative of the fixation of high performance alleles and/or the optimization of allelic arrays (i.e., epistatic complementation) after one cycle of selection (Robbins et al. 2008). Lack of response to selection might also be partially attributable to undetectable genetic gains due to substantial environmental factors (environmental variance >50%; Table 2) that affect yield-related traits in cucumber (e.g., sex expression and lateral branch number) (Serquen et al. 1997a; Fazio et al. 2003b; Fan et al. 2006).

The initial response to selection (F_3–F_4; PHE + MAS) for all of the traits examined (Table 2) was similar to that yield components reported by Fan et al. (2006) during backcrossing of Gy-7 × H-19-derived germplasm. In our recurrent selection study, the genotypic variances (σ^2G) associated with yield, L:D and E:T increased from 20.63, 0.15 and 0.002 in F_3–44.06, 0.26 and 0.003 in F_4 due to selection (PHE + MAS) (Table 2). In contrast, the environmental variances (σ^2E) associated with yield, L:D and E:T decreased from 73.06, 0.93, and 0.002 in F_3–53.51, 0.55 and 0.001 in F_4 due to selection (PHE + MAS).

A greater positive response to PHE + MAS selection (Table 2) might have been predicted if the traits under selection were conditioned by few additive genes (perhaps 2–3). Gynoecy in cucumber is influenced by at least five modifying genes (Serquen et al. 1997a, b; Fazio et al. 2003b) that interact epistatically and are environmentally dependent (Serquen et al. 1997b; Fazio 2001). Similarly, MLB is controlled by at least four major genes with additive epistatic effects (Serquen et al. 1997a; Fazio et al. 2003a). A decrease in the frequencies of yield-related alleles (MLB) at the three marker loci (AK5SCAR, BC523SCAR and CSWTAAA01) was detected after initial (F_3–F_4) selection (Table 4).

These QTL are located on different linkage groups (AK5SCAR on LG 6; CSWTAAA01 on LG 4; BC523SCAR on LG 1) (Fazio et al. 2003b). Response to MAS for MLB is dependent on the magnitude of QTL effects, the distance between that marker and a specific QTL, and the environmental variance (Fazio et al. 2003b). Fan et al. (2006) applied five marker loci for improvement of gynoecy that increased by 5.6–9.8% per cycle by MAS. However, ΔG of MLB and L:D was not sufficient to contribute to an increase in yield per se. Our results and those of others (Fazio et al. 2003a, b; Robbins et al. 2008) indicate that yield attributes of cucumber are dependent on the effects of several QTL that are affected by epistatic and environmental interactions.

Several factors are important for successful implementation MAS for plant improvement. Marker efficiency can be potentiated when marker-trait associations are unique (LOD > 3.0 and $R^2 > 5\%$) and positioned on a high-density (mean marker interval is <3 cM) genetic map (Staub et al. 1996). Furthermore, appropriate population size and marker choice and number are primary considerations for maximizing response to MAS for traits having low heritabilities (Moreau et al. 1998). Some of the markers employed herein have comparatively loose marker/trait linkage associations (Fazio et al. 2003a), which likely compromised significant ΔG through MAS. Yield increases in the populations described herein might have been enhanced if additional markers and progeny were used in MAS, and evaluation environments were expanded. For instance, the positive allelic effect associated with CSWTAA0B and OPAT15-1 (i.e., QTL conditioning with E:T ratio) was fixed in advanced generations (F_4 and F_5) after PHE + MAS (Table 4). These results indicate that selection for alleles with positive effects at these maker loci did not result in a positive ΔG for E:T. Nevertheless, even though greater positive response to selection for yield and quality traits examined herein might have been attained using additional linked markers having additive effects, pyramiding of such loci can be unpredictable and knowledge of their epistatic interactions and source/sink relationships must be known (Robbins et al. 2008). Without such information, response from MAS may not be highly predictable.

The deployment of MAS has been effective in increasing ΔG in cucumber for yield components during single (Fazio et al. 2003b) and multiple (Fan et al. 2006) trait backcrossing, but less effective during recurrent selection during population development (Robbins and Staub 2009). The ΔG after the initial cycle of MAS reported herein, however, was not remarkable, and mirrors other finding in a diverse array of crop species (Gimelfarb and Lande 1994; Hospital et al. 1997; Moreau et al. 2000). A promising breeding strategy for MAS was described by Hospital et al. (1997), where the first cycle of combined selection (MAS + PHE) was followed in subsequent generations genotyping using markers tightly linked to the target trait. This selection strategy was found to be effective even for traits possessing low to moderate heritability (Moreau et al. 2000). Improvement of traits of low heritability in cucumber (e.g., yield and quality components) that are associated with QTL having complicated negative associations and epistatic effects may benefit from the application of such selection strategies.

Acknowledgments The fund provided by the Department of Biotechnology, Ministry of Science and Technology, Govt. of India for sponsoring T.K. Behera to carry out research work at Department of Horticulture, UW Madison, USA is highly acknowledged.

References

Behera TK, Staub JE, Behera S, Rao AR, Mason S (2008) One cycle of phenotypic selection combined with marker assisted selection for improving yield and quality in cucumber. In: Pitrat M (ed) Proceedings of the IXth EUCARPIA meeting on genetics and breeding of Cucurbitaceae. Avignon, France, pp 115–121

Dudley JW (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Sci 33:660–668

Edwards MD, Stuber CW, Wendel JF (1987) Molecular marker-facilitated investigations of quantitative trait loci in maize. I. Numbers, distribution, and types of gene action. Genetics 116:113–125

Fan Z, Robbins MD, Staub JE (2006) Population development by phenotypic selection with subsequent marker-assisted selection for line extraction in cucumber (*Cucumis sativus* L.). Theor Appl Genet 112:843–855

Fazio G (2001) Comparative study of marker assisted and phenotypic selection and genetic analysis of yield components in cucumber. PhD dissertation, University of Wisconsin, Madison

Fazio G, Staub JE, Stevens MR (2003a) Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Theor Appl Genet 107:864–874

Fazio G, Chung SM, Staub JE (2003b) Comparative analysis of response to phenotypic and marker-assisted selection for

multiple lateral branching in cucumber (*Cucumis sativus* L.). Theor Appl Genet 107:875–883

Flint-Garcia SA, Darrah LL, McMullen MD, Hibbard BE (2003) Phenotypic versus marker-assisted selection for stalk strength and second-generation European corn borer resistance in maize. Theor Appl Genet 107:1331–1336

Gimelfarb A, Lande R (1994) Simulation of marker assisted selection in hybrid populations. Genet Res 63:39–47

Hospital F, Moreau L, Lacoudre F, Charcosset A, Gallais A (1997) More on the efficiency of marker-assisted selection. Theor Appl Genet 95:1181–1189

Kasha KJ (1999) Biotechnology and world food supply. Genome 42:642–645

Kennard WC, Havey MJ (1995) Quantitative trait analysis of fruit quality in cucumber: QTL detection, confirmation, and comparison with mating design. Theor Appl Genet 91:53–61

Knapp SJ (1998) Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. Crop Sci 38:1164–1174

Lande R, Thompson R (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743–756

Lecomte L, Duffe P, Buret M, Servin B, Hospital F, Causse M (2004) Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. Theor Appl Genet 109:658–668

Littell RC, Milliken GA, Stroup WW, Wolfinger RD (1996) SAS system for mixed models. SAS Institute Inc, Cary

Lower RL, Edwards MD (1986) Cucumber breeding. In: Basset MJ (ed) Breeding vegetable crops. AVI Publishing Co., Westport, pp 173–207

Miller MP (1997) Tools for population genetic analysis (TEPGA), Version 3. Department of Biological Sciences. Northern Arizona University, Arizona, USA

Moreau L, Charcosset A, Hospital F, Gallais A (1998) Marker assisted selection efficiency in populations of finite size. Genetics 148:1353–1365

Moreau L, Lamarie S, Charcosset A, Gallais A (2000) Economic efficiency of one cycle of marker assisted selection. Crop Sci 40:329–337

Moreau L, Charcosset A, Gallais A (2004) Experimental evaluation of several cycles of marker-assisted selection in maize. Euphytica 137:111–118

Nei M (1972) Genetic distance between populations. Am Nat 106:283–292

Ortiz R (1998) Critical role of plant biotechnology for the genetic improvement of food crops: Perspectives for the next millennium. Electron J Biotechnol 1:152–159

Paterson AH, Tanksley SD, Sorrells ME (1991) DNA markers in plant improvement. Adv Agron 46:39–90

Robbins MD (2006) Molecular marker development, QTL pyramiding, and comparative analysis of phenotypic and marker assisted selection in cucumber. PhD dissertation University of Wisconsin, Madison

Robbins MD, Staub JE (2009) Comparative analysis of marker-assisted and phenotypic selection for yield components in cucumber. Theor Appl Genet 119:621–634

Robbins MD, Casler M, Staub JE (2008) Pyramiding QTL for multiple lateral branching in cucumber using nearly isogenic lines. Mol Breed 22:131–139

SAS software, Version 8 of SAS System for Unix (1999) SAS Institute Inc., Cary

Serquen FC, Bacher J, Staub JE (1997a) Genetic analysis of yield components in cucumber (*Cucumis sativus* L.) at low plant density. J Am Soc Hortic Sci 122:522–528

Serquen FC, Bacher J, Staub JE (1997b) Mapping and QTL analysis of a narrow cross in cucumber (*Cucumis sativus* L.) using random amplified polymorphic DNA markers. Mol Breed 3:257–268

Staub JE, Bacher J (1997) Cucumber as a processed vegetable. In: Smith DS, Cash JN, Nip W, Hui YH (eds) Processing vegetables: science and technology IV. Technomic Publishing Co, Inc Lancaster, PA, pp 129–193

Staub JE, Grumet R (1993) Selection for multiple disease resistance affects cucumber yield potential. Euphytica 67:205–213

Staub JE, Balgooyen B, Tolla GE (1986) Quality and yield of cucumber hybrids using gynoecious and bisexual pollen parents. HortScience 21:510–512

Staub JE, Serquen FC, Gupta M (1996) Genetic markers, map construction and their application in plant breeding. HortScience 31:729–741

Steele R, Torrie JH (1980) Principles and procedures of statistics, 2nd edn. McGraw-Hill, New York

Steele KA, Edwards G, Zhu J, Witcombe RC (2004) Marker evaluated selection in rice: shifts in allele frequency among bulks selected in contrasting agricultural environments identify genomic regions of importance to rice adaptation and breeding. Theor Appl Genet 109:1247–1260

Tanksley SD, Medino-filho H, Rick CM (1981) The effect of isozyme selection on metric characters in an inter-specific backcross of tomato: Basis of an early screen procedure. Theor Appl Genet 60:291–296

Tatlioglu T (1993) Cucumber (*Cucumis sativus* L.). In: Kalloo GB, Bergh O (eds) Genetic improvement of vegetable crops. Pergamon Press Ltd, Tarrytown, pp 197–234

Thabuis A, Palloix A, Servin B, Daubeze AM, Signoret P, Hospital F, Lefebvre V (2004) Marker-assisted introgression of 4 *Phytophthora capsici* resistance QTL alleles into a bell pepper line: validation of additive and epistatic effects. Mol Breed 14:9–20

Wehner TC (1989) Breeding for improved yield in cucumber. Plant Breed Rev 6:323–359

Willcox MC, Khairallah MM, Bergvinson D, Crossa J, Deutsch JA, Edmeades GO, Gonzalez de Leon D, Jiang C, Jewell DC, Mihm JA, Williams WP, Hoisington D (2002) Selection for resistance to southwestern corn borer using marker-assisted and conventional backcrossing. Crop Sci 42:1516–1528

Xie C, Xu S (1998) Efficiency of multistage marker-assisted selection in the improvement of multiple quantitative traits. Heredity 80:489–498

Yeh FC, Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belg J Bot 129:157